

Score-Based Generative Modeling with Critically-Damped Langevin Diffusion

Year: 2021 | Citations: 267 | Authors: Tim Dockhorn, Arash Vahdat, Karsten Kreis

Abstract

Score-based generative models (SGMs) have demonstrated remarkable synthesis quality. SGMs rely on a diffusion process that gradually perturbs the data towards a tractable distribution, while the generative model learns to denoise. The complexity of this denoising task is, apart from the data distribution itself, uniquely determined by the diffusion process. We argue that current SGMs employ overly simplistic diffusions, leading to unnecessarily complex denoising processes, which limit generative modeling performance. Based on connections to statistical mechanics, we propose a novel critically-damped Langevin diffusion (CLD) and show that CLD-based SGMs achieve superior performance. CLD can be interpreted as running a joint diffusion in an extended space, where the auxiliary variables can be considered “velocities” that are coupled to the data variables as in Hamiltonian dynamics. We derive a novel score matching objective for CLD and show that the model only needs to learn the score function of the conditional distribution of the velocity given data, an easier task than learning scores of the data directly. We also derive a new sampling scheme for efficient synthesis from CLD-based diffusion models. We find that CLD outperforms previous SGMs in synthesis quality for similar network architectures and sampling compute budgets. We show that our novel sampler for CLD significantly outperforms solvers such as Euler–Maruyama. Our framework provides new insights into score-based denoising diffusion models and can be readily used for high-resolution image synthesis. VPSDE. We leave the study of CLD with maximum likelihood training for high-dimensional (image) datasets to future work.